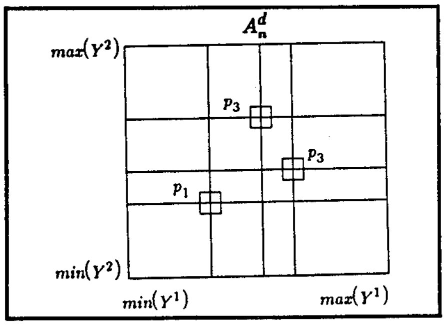
Fast Multidimensional Density Estimation based on Random-width Bins

Leonard B. Hearne
and
Edward J. Wegman

Technical Report No. 109
October, 1994

Center for Computational Statistics



George Mason University
Fairfax, VA 22030

CENTER FOR COMPUTATIONAL STATISTICS
TECHNICAL REPORT SERIES (RECENT REPORTS)

TR 93. Winston C. Chow, Modeling and Estimation with Fractional Brownian Motion and Fractional Gaussian Noise (Ph.D. Dissertation), February, 1994.

TR 94. Mark C. Sullivan and Edward J. Wegman, Correlation Estimators Based on Simple Nonlinear Transformations, February, 1994, To appear **IEEE Transactions on Signal Processing**.

TR 95. Mark C. Sullivan and Edward J. Wegman, Normalized Correlation Estimators Based on Simple Nonlinear Transformations, March, 1994.

TR 96. Kathleen Perez-Lopez and Arun Sood, Comparison of Subband Features for Automatic Indexing of Scientific Image Databases, March, 1994.

TR 97. Wendy L. Poston and Jeffrey L. Solka, A Parallel Method to Maximize the Fisher Information Matrix, June, 1994.

TR 98. Edward J. Wegman and Charles A. Jones, Simulating a Multi-target Acoustic Array on the Intel Paragon, June, 1994.

TR 99. Barnabas Takacs, Edward J. Wegman and Harry Wechsler, Parallel Simulation of an Active Vision Model, June, 1994.

TR 100. Edward J. Wegman and Qiang Luo, Visualizing Densities, October, 1994.

TR 101. Daniel B. Carr, Converting Tables to Plots, October, 1994.

TR 102. Julia Corbin Fauntleroy and Edward J. Wegman, Parallelizing Locally-Weighted Regression, October, 1994.

TR 103. Daniel B. Carr, Color Perception, the Importance of Gray and Residuals on a Choropleth Map, October, 1994.

TR 104. David J. Marchette, Carey E. Priebe, George W. Rogers and Jeffrey L. Solka, Filtered Kernel Density Estimation, October, 1994.

TR 105. Jeffrey L. Solka, Edward J. Wegman, Carey E. Priebe, Wendy L. Poston and George W. Rogers, A Method to Determine the Structure of an Unknown Mixture Using the Akaike Information Criterion and the Bootstrap, October, 1994.

TR 106. Wendy L. Poston, Edward J. Wegman, Carey E. Priebe and Jeffrey L. Solka, A Contribution to the Theory of Robust Estimation of Multivariate Location and Shape: EID, October, 1994.

TR 107. Clifton D. Sutton, Tree Structured Density Estimation, October, 1994.

TR 108. Charles A. Jones, Simulating a Multi-target Acoustic Array on the Intel Paragon (M.S. Thesis), October, 1994.

TR 109. Leonard B. Hearne and Edward J. Wegman, Fast Multidimensional Density Estimation based on Random-width Bins, October, 1994.

Fast Multidimensional Density Estimation based on Random-width Bins

Leonard B. Hearne[1] and Edward J. Wegman[2]
Center for Computational Statistics
George Mason University
Fairfax, VA 22030

Abstract

Histogram-type density estimators have some notable computational advantages over other forms of density estimation by virtue of the WARPing algorithm. However, traditional fixed-bin-width have less than satisfactory smoothing properties, being too coarse in regions of high density and too fine in regions of low density. Scott (1992) suggests the ASH algorithm as a means of overcoming these problems, but the ASH algorithm is computationally intensive somewhat negating the benefits of WARPing. Wegman (1975) proposed a variable bin-width technique for one dimensional density estimators and used sieve-type methods to show strong consistency results that did not depend on smoothness properties of the underlying density. In this paper, we extend this idea to high-dimensional, variable bin-width meshes. The boundaries of the bins are determined by a random subsampling of the observations. An extension of the WARPing algorithm may still be used for fast computation. We give combinatorial arguments for calculating the number of bins and also the conditional expectation and variance of the number of observations per bin. Conditional on the random hyper-rectangular tessellation, we calculate the maximum likelihood density estimator.

Introduction

In this paper, a density estimation method is developed that is computationally more tractable than kernel density methods, and has better smoothing properties than traditional fixed binning methods. The basic method is easy to describe in one dimension. Randomly select a subset of m observations $\left\{Y^*\right\}$ from a set of n observations $\{Y\}$, $m < n$, together with the $max\{Y\}$ and $min\{Y\}$. Order the set $\left\{Y^*\right\}$ in the set $\left\{Y^*_{(\cdot)}\right\}$. A set of random width bins $\{B\}$ can be can be constructed using adjacent elements in the set $\left\{Y^*_{(\cdot)}\right\}$. Then attribute the probability mass of all observations in $\{Y\}$ to the bins in $\{B\}$. The probability density on an element $B_i \in \{B\}$ is the relative probability mass on B_i divided by the length of B_i, *cf.* Wegman (1975) and Hearne and Wegman (1991). There are many ways to generalize these results to a d-dimensional support space. The generalization that we have adopted here is to define random-width d-dimensional rectangular bins generated by a random sample from the set of observations.

Random-width d-Dimensional Bin Tessellation

[1] The research of Dr. Hearne was supported by the National Security Agency under contract MDA 904-92-H-3082

[2] The research of Dr. Wegman was supported by the Army Research Office under Contract DAAL03-91-G-0039, the Office of Naval Research under Grant N00014-92-J-1303, and the National Science Foundation under Grant DMS9002237.

Given a set of n observations, $\{Y\}$, in a d-dimensional Euclidian space, let A_n^d be the minimum d-dimensional rectangular cover of $\{Y\}$. Each observation $Y_j \in \{Y\}$ can be written in the form $Y_j = \left(Y_j^1, Y_j^2, \cdots, Y_j^d\right)$. Then A_n^d can be defined by the set of maximum and minimum values for the d coordinate axes,

$$A_n^d \equiv \left\{ \boldsymbol{x} \in \Re^d \colon \boldsymbol{x}^i \geq min\left(Y^i\right) \wedge \boldsymbol{x}^i \leq max\left(Y^i\right) \right\}.$$

A d-dimensional rectangular tessellation of A_n^d can be generated by selecting a random subsample of N observations $\left\{S_N\right\}$ from $\{Y\}$. For each of the d coordinate axes let $\left\{S_N^i\right\}$ be the set of the i^{th} coordinate for all $Y \in \left\{S_N\right\}$ together with $max\left(Y^i\right)$ and $min\left(Y^i\right)$. Let $\left\{S_{(\cdot)}^i\right\}$ be the ordered set of unique elements in $\left\{S_N^i\right\}$ and $s^i = card\left\{S_{(\cdot)}^i\right\}$. A set of one dimensional bins, $\left\{B^i\right\}$, can be generated for each of the d coordinate axes by adjacent elements in the set $\left\{S_{(\cdot)}^i\right\}$, and $card\left\{B^i\right\} = s^i - 1$. The d-dimensional rectangular random tessellation $\left\{B_N^d\right\}$ of A_n^d can then be generated by the cross product of the sets of one dimensional bins for each coordinate axis;

$$\left\{B_n^d\right\} = \left\{B^1\right\} \times \left\{B^2\right\} \times \cdots \times \left\{B^d\right\}, \text{ and}$$

$$m = card\left\{B_n^d\right\} = \prod_{i=1}^{d} \left(s^i - 1\right).$$

The upper bound on the cardinality of the set of one dimensional bins that are generated for each of the coordinate axes is $s^i - 1 \leq N + 1$, $1 \leq i \leq d$, since the random sample $\left\{S_N\right\}$ may have observations that contain $max\left(Y^i\right)$ or $min\left(Y^i\right)$, observations are recorded only to finite precision, and computers operate on a subset to the rational numbers. The cardinality of the tessellation $\left\{B_N^d\right\}$ then has an upper bound, given the random subsample $\left\{S_N\right\}$ of

$$m = card\left\{B_n^d\right\} = \prod_{i=1}^{d} \left(s^i - 1\right) \leq \left(N + 1\right)^d.$$

In Figure 1 a set of observations $\{Y\}$ in $\Re^2$ have values $max\left(Y^1\right)$, $min\left(Y^1\right)$, $max\left(Y^2\right)$, and $min\left(Y^2\right)$. These values define the minimum 2-dimensional rectangular cover A_n^2 of $\{Y\}$. A random subsample of observations is drawn from $\{Y\}$, $\left\{S_3\right\} \equiv \left(p_1, p_2, p_3\right)$. These three points together with the maximum and minimum values for each of the coordinate axes generate the set of bins $\left\{B_n^2\right\}$ of A_n^2.

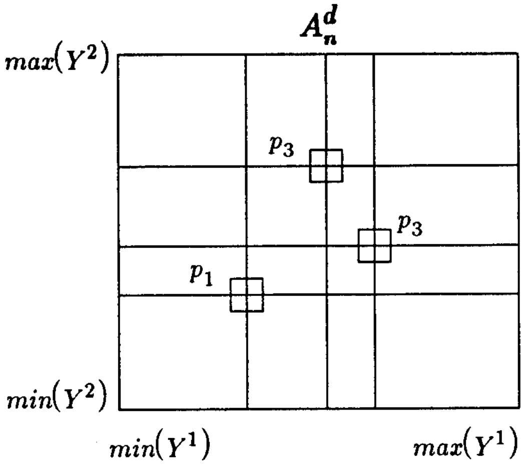


Figure 1

The tessellation $\left\{B_n^d\right\}$ of A_n^d is adaptive in the sense that the elements of the tessellation tend to be large where the observations are sparse and small where the observations are not sparse.

Conditional Expectation and Variance of the Number of Observations per Bin

Let B_k, $1 \leq k \leq m$, be the k^{th} d-dimensional bin in the tessellation $\left\{B_n^d\right\}$ of A_n^d, and let Z_k be the number of observations in $\{Y\}$ that are in B_k. The expected value of Z_k given the tessellation $\left\{B_n^d\right\}$ is the number of observations that might be attributed to the k^{th} bin times the probability that the d-dimensional random variable X is in the k^{th} bin;

$$E\left[Z_k \mid \left\{B_n^d\right\}\right] = (n - N)P\left(X \in B_k\right).$$

Let U^i_j, $1 \le i \le d$, be the empirical probability mass on the j^{th} one dimensional bin, $1 \le j \le s^i - 1$, for the i^{th} coordinate axis,

$$U^i_j = F\left(Y^i_{(j-1)}\right) - F\left(Y^i_{(j)}\right) = P\left(X^i \in B^i_j \mid \{B^i\}\right).$$

Using order statistical arguments, *cf.* Rohatgi (1976) pp.575-580, it can be shown that;

$$E\left[U^i_j \mid \{B^i\}\right] = \frac{1}{s^i - 1},\ 1 \le j \le s^i - 1, \text{ and}$$

$$V\left[U^i_j \mid \{B^i\}\right] = \frac{s^i - 2}{\left(s^i - 1\right)^2 s^i}.$$

Since the tessellation $\left\{B^d_n\right\}$ of A^d_n is generated by the cross product of the one dimensional bins on each of the d coordinate axes then the probability mass that is on a given d-dimensional bin $B_k \in \left\{B^d_n\right\}$, given the tessellation $\left\{B^d_n\right\}$, is;

$$E\left[U_k \mid \left\{B^d_n\right\}\right] = \prod_{i=1}^{d} \frac{1}{s^i - 1},\ 1 \le k \le m, \text{ and}$$

$$V\left[U_k \mid \left\{B^d_n\right\}\right] = \prod_{i=1}^{d} \frac{s^i - 2}{\left(s^i - 1\right)^2 s^i}.$$

Multiplying by the number of observations that might be attributed to a d-dimensional rectangular bin, $n - N$, and applying the inequality bounding the cardinality of the number of bins in the tessellation;

$$E\left[Z_k \mid \left\{B^d_n\right\}\right] \ge \frac{n - N}{\left(N + 1\right)^d},\ 1 \le k \le m, \text{ and}$$

$$V\left[Z_k \mid \left\{B^d_n\right\}\right] \ge \frac{(n - N)^2 (N - 1)^d}{N^{2d}\left(N + 1\right)^d}.$$

A Class of Probability Density Estimators

Let n be the number of observations in the set of observations $\{Y\}$, and let n_k be the number of observations in the k^{th} rectangular bin in the tessellation $\left\{B^d_n\right\}$. Let $W\left(N_k\right)$ be the probabilistic mass of observations in the tessellation generating set $\left\{S_N\right\}$ that are attributed to an adjacent bin in the tessellation $B_k \in \left\{B^d_n\right\}$ by the function $W(\cdot)$. And let C_k be the d-dimensional content of the k^{th} element of the tessellation. Then we can define a class of probability density estimators on a tessellation $\left\{B^d_n\right\}$ by;

$$\hat{f}\left(x \in B_k\right) = \frac{n_k + W\left(N_k\right)}{n \cdot C_k} \text{ and}$$

$$\hat{f}\left(x \notin \left\{B^d_n\right\}\right) = 0.$$

This class of probability density estimators is constant on each bin in the tessellation, and the content of each of the d-dimensional bins in the tessellation C_k is easily computed. The probabilistic mass attribution function $W(\cdot)$ is closely related to the likelihood function.

The Likelihood Function

The likelihood function was introduced as a means for optimizing the parameter values in the parametric density estimation setting so that the fitted parametric function would best fit a set of observations. In the nonparametric setting the likelihood function has utility if there is a variable in the class of density estimators. The weight that is attributed to bins in the tessellation by observations in $\left\{S_N\right\}$ is variable and can be used to optimize the likelihood function.

The likelihood function for this class of probability density estimators is

$$L(x) = \prod_{j=1}^{n} \frac{n_k + W\left(N_k\right)}{n \cdot C_k},$$

the product of the density estimates for each of the observations. But the class of density estimators that are presented here are estimators on the set of bins in the tessellation of A^d_n so the likelihood function can be reformulated in terms of the elements of the tessellation;

$$L(\boldsymbol{x}) = \prod_{k=1}^{m} \left(\frac{n_k + W(N_k)}{n \cdot C_k} \right)^{(n_k + W(N_k))}.$$

Taking the first derivative of the log of the likelihood function with respect to $W(N_k)$;

$$\frac{d}{dW(N_k)} log L(\boldsymbol{x}) = \sum_{k=1}^{m} \left(\frac{n_k}{n_k + W(N_k)} + \frac{W(N_k)}{n_k + W(N_k)} \right)$$
$$+ \sum_{k=1}^{m} \left(log(n_k + W(N_k)) - log(n \cdot C_k) \right).$$

If the first derivative is set equal to zero and solved for $W(N_k)$ then the estimator will be optimized, either maximized or minimized depending on the sign of the second derivative of the log of the likelihood function. Taking the second derivative of the log of the likelihood function;

$$\frac{d^2}{dW(N_k)^2} log L(\boldsymbol{x}) = \sum_{k=1}^{m} \frac{n_k}{n_k + W(N_k)}.$$

The second derivative of the log of the likelihood function with respect to $W(N_k)$ is positive on all bins in the tessellation that have observations in them, $n_k > 0$, and is undefined where $n_k = 0$. The likelihood function is thus convex and the likelihood function is maximized when the probabilistic mass of all observations in $\{S_N\}$ are attributed to the adjacent bin where $\frac{n_k + 1}{C_k}$ will be largest.

A Random Bin-width Warping Algorithm

For the proposed probability density estimation method to be of utility it is important that density estimates be readily computable, given a set of n observations, $\{Y\}$, in a d-dimensional Euclidian space. The principal computational complexity is in the attribution of observations to bins in the tessellation, $\{B_n^d\}$, of the minimum d-dimensional rectangular cover of $\{Y\}$, A_n^d. In conventional fixed width binning methods an algorithm called warping has been developed that increases the speed and reduces the computational complexity for attributing observations to bins in the tessellation. This algorithm has been extended to variable bin-width tessellations.

Given N the number of observations in the random sample of observations used to generate the rectangular bins in the tessellation, the cardinality of the set of bins, m, is bounded by;

$$m = card\{B_n^d\} = \prod_{i=1}^{d} (s^i - 1) \leq (N + 1)^d.$$

For each coordinate axis there is an upper bound on the number of one dimensional bins that can be generated. Let Bound_Values[i, j] be a matrix with the i^{th} row, $0 \leq i < d$, corresponding to $\{S_{(\cdot)}^i\}$ and Bound_Value[i, 0]=$min(Y^i)$. Then for each row i, $0 \leq j \leq s^i - 1$. Let Bin_Index[i, k] be a matrix with the i^{th} row a vector of integer indices into the matrix Bound_Values[i, j], with $0 \leq k < w^i$, where w^i is the selected number of warping indices for the i^{th} coordinate axis, $s^i - 1 \leq w^i$.

Let $b^i = min(Y^i)$ and $a^i = \frac{max(Y^i) - min(Y^i)}{w^i}$ for the i^{th} coordinate axis, $0 \leq i < d$. For any point $x^i \in [min(Y^i), max(Y^i)]$ then the value

$$\text{Index} = \text{Truncate}\left[\frac{(x^i - b^i)}{a^i}\right]$$

is an integer in the range $0 \leq \text{Index} < w^i$. Let the i^{th} coordinate axis and the k^{th} entry in the matrix Bin_Index[i, k] be the smallest index j into the matrix Bound_Values[i, j] such that

$$a^i(\text{Index} + b^i) \leq \text{Bound\_Values}[i, j].$$

Then an efficient algorithm to compute the bin index for the i^{th} coordinate axis, $0 \leq i < d$, is shown in the following code fragment.

```
Get_Bin_Index($i, x^i$)
    Table_Index = Truncate$\left((x^i - b^i)/a^i\right)$
    Index = Bin_Index[$i$, Table_Index]
    While($x^i >$ Bound_Values[$i$, Index]) Index++
    Return Index
```

The size of the number of warping indices, w^i, is specified by the user of the density estimation method. The question of how large w^i should be is of interest. We want to maximize the probability of selecting the correct bin index on the first attempt for each of the d coordinate axes. The bounds on the probability of selecting the correct bin index on the first attempt is;

$$P\left(x^i < \text{Bound\_Values}[i,\text{Bin\_Index}[i,\text{Table\_Index}]]\right) \geq \frac{w^i - s^i + 1}{w^i}.$$

The larger w^i is relative to $s^i - 1$, the larger the probability that the correct bin index will be computed on the first attempt. If the density function is symmetric then the expected value of the probability is $\frac{w^i - (s^i - 1)/2}{w^i}$.

Conclusions and Extensions

Random-width binning methods are a computationally tractable alternative to fixed-width binning methods. The size of the bins in a d-dimensional space are adaptive so that the bins will tend to be large where the observations are sparse and small where the observations are not sparse. Bounds on the expected value and variance of the number of observations that are attributed to each bin can be calculated, given the size of the subsample that is randomly selected from the set of observations to generate the d-dimensional bins. The likelihood function is convex a function that can be maximized or minimize to give a maximum entropy estimate by selecting the appropriate probabilistic weight distribution function $W(\cdot)$, *cf.* Hearne and Wegman (1992). By applying an extension to the WARPing algorithm, the computational complexity of the random-width binning method is only slightly more computationally intensive than fixed-width binning methods.

One of the natural extensions to random-width binning methods is to apply a resampling scheme, *cf.* Billard and LaPage (1992). Given smoothness assumptions about the underlying probability density, then the size of the set of observations, the dimension of the observations space, and the expected value and variance bound on the number of observations that are attributed to each bin might be used to find the optimal subsample size, and the number of resampling repetitions necessary to achieve the desired density estimate smoothness. Resampling in an optimal way is believed to be less computationally intensive than either kernel or ASH methods, *cf.* Scott (1992).

Bibliography

Hearne, L.B. and Wegman E.J. (1991). "Adaptive Probability Density Estimation in Lower Dimensions using Random Tessellations", Computing Science and Statistics, Keramidas, E.M. (ed.), 23 241-245, Interface Foundation of North America, Fairfax Station, VA.

Hearne, L.B. and Wegman E.J. (1992). "Maximum Entropy Density Estimation using Random Tessellations", Computing Science and Statistics, Newton J. (ed.), **24** 483-487, Interface Foundation of North America, Fairfax Station, VA.

LePage, R. and Billard, L. (1992). Exploring the Limits of Bootstrap. John Wiley & Sons, New York.

Rohatgi, V.K. (1976). An Introduction to Probability Theory and Mathematical Statistics. John Wiley & Sons, New York.

Wegman, E.J. (1975). "Maximum Likelihood Estimation of a Probability Density Function" Sankhyā Ser. A **37** 211-224.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
OCTOBER, 1994

3. REPORT TYPE AND DATES COVERED
TECHNICAL

4. TITLE AND SUBTITLE
Fast-Multidimentional Density Estimation Based on Random-width Bins

5. FUNDING NUMBERS
DAAL03-91-G-0039
DAAH04-94-G-0267

6. AUTHOR(S)
Leonard B. Hearne and Edward J. Wegman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Computational Statistics
George Mason University
Fairfax, VA 22030

8. PERFORMING ORGANIZATION REPORT NUMBER
TR no. 109

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Histogram-type density estimators have some notable computational advantages over other forms of density estimation by virtue of the WARPing algorithm. However, traditional fixed -bin-width have less than satisfacory smoothing properties, being too coarse in regions of high density and too fine in regions of low density. Scott (1992) suggests the ASH algorithm as a means of overcoming these problems, but the ASH algorithm is computationally intensive somewhat negating the benefits of WARPing. Wegman (1975) proposed a variable bin-width technique for one demensional density estimators and used sieve-type methods to show strong consisency results that did not depend on smoothness properties of the underlying density. In this paper, we extend this idea to high-demensional, variable bin-width meshes. The boundaries of the bins are determined a a random subsampling of the observations. An extension of the WARPing algorithm may still be used for fast computation. We give combinatorial arguements for calculating the number of bins and also the conditional expectation and varianceof the number of observations per bin. Conditional on the random hyper-rectangular tessellation, we calculate the maximum likelihood density estimator.

14. SUBJECT TERMS
ASH, WARPing, random bins, random tessellations, maximum likelihood

15. NUMBER OF PAGES
8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
October, 1994

3. REPORT TYPE AND DATES COVERED
Technical

4. TITLE AND SUBTITLE
Fast-Multidemtional Density Estimation Based on
Random-width Bins

5. FUNDING NUMBERS
N00014-92-J-1303
N00014-93-1-0527

6. AUTHOR(S)
Leonard B. Hearne and Edward J. Wegman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Center for Computational Statistics
George Mason University
Fairfax, VA 22030

8. PERFORMING ORGANIZATION REPORT NUMBER
TR no. 109

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Department of the Navy
Office of the Chief of Naval Research
Mathematical Sciences Division
800 N. Quincy Street Code 1111SP
Arlington, VA 22217-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Navy position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Histogram-type density estimators have some notable computational advantages over other forms of density estimation by virtue of the WARPing algorithm. However, traditional fixed -bin-width have less than satisfacory smoothing properties, being too coarse in regions of high density and too fine in regions of low density. Scott (1992) suggests the ASH algorithm as a means of overcoming these problems, but the ASH algorithm is computationally intensive somewhat negating the benefits of WARPing. Wegman (1975) proposed a variable bin-width technique for one demensional density estimators and used sieve-type methods to show strong consisency results that did not depend on smoothness properties of the underlying density. In this paper, we extend this idea to high-demensional, variable bin-width meshes. The boundaries of the bins are determined a a random subsampling of the observations. An extension of the WARPing algorithm may still be used for fast computation. We give combinatorial arguements for calculating the number of bins and also the conditional expectation and varianceof the number of observations per bin. Conditional on the random hyper-rectangular tessellation, we calculate the maximum likelihood density estimator.

14. SUBJECT TERMS
ASH, WARPing, random bins, random tessellations, maximum likelihood

15. NUMBER OF PAGES
8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL